

Association for Information Systems

AIS Electronic Library (AISeL)

ICEB 2002 Proceedings

International Conference on Electronic Business
(ICEB)

Winter 12-10-2002

English as the Infrastructure Language for a Multilingual Internet

Garry White

Follow this and additional works at: <https://aisel.aisnet.org/iceb2002>

This material is brought to you by the International Conference on Electronic Business (ICEB) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ICEB 2002 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.

English as the Infrastructure Language for a Multilingual Internet

Garry White

Computer Information Systems Department
Southwest Texas State University
San Marcos, Texas U.S.A.
gw06@business.swt.edu

Abstract

English has been the dominant language of the Internet for users and infrastructure development. This is changing due to the increasing number of non-English speaking Internet users. While the Internet community will eventually become a multilingual environment; the development of the Internet infrastructure source will still be English. This will have an impact on global businesses and government's education systems in non-English speaking countries. Eventually, English will be the global language for Internet developers and professionals of global businesses.

1. Introduction

In 1999, less than 10% of the world's populations spoke English as their primary language. The majority (56.5%) of Web users use English as the language of choice [4]. This is changing. Global Reach has indicated that one half of Internet users are from non-English speaking nations. By 2003, this will reach 70%. The majority of Web page content will be in languages other than English [11]. There is more and more evidence of the adoption of non-English use of the Internet. Evidence of this growing trend are the emergence of translation services, multilingual domain names, Unicode, and the use of icons [17]. It is clear that Web pages are becoming multilingual for users and business customers. However, the infrastructure and Internet development are still in English. This is due to the need for global standards in programming and technical communication.

2. The English Internet

The Internet began with ARPANET, a research project by the U.S. Department of Defense, to allow military computers to communicate through dynamic paths of communications. If one path failed, another path was used. Later the U.S. government divided this network into a military subsystem and a civilian subsystem. It was the civilian network that became today's Internet. In 1992, the U.S. government allowed businesses to link to the Internet. [16, p. 243]

The character set used by this new network was American Standard Code for Information Interchange (ASCII). This has been the coding scheme for computer communication. Like the Internet, it too originated in the United States. The text characters used consisted of the English alphabet. E-mail and web page source code use

this character set.

As expected, since the development of the Internet originated in an English speaking country, the technical infrastructure and web pages (interface for users) all have used English. English has been the lingua franca of the Internet [15].

3. Need for Multilingual Web Pages for Customers

The Internet has created one global market. Those who understand the technical and cultural issues have a head start on capturing new audiences and helping their businesses forge a worldwide market [22]. There is a need for a multilingual Web. By having Web sites in local languages, more businesses and opportunities can be established. People are more comfortable with e-commerce when the product is in their native language [24], even if English is their second language. Those companies that only have English version of their Web sites or just their local language versions stand to lose their existing or potential international clients to their competitors who communicate in the native language of the customer [17].

Several Information Technology trade periodicals indicate that English has been the dominant language on the Internet and most of the revenue generated inside the U.S.A. However, this is changing. The Internet is moving from English to multilingual [7][5][14][20][22][25][26][29]. For example, English is a second language in Thailand. However, Thai language content web is increasing [6]. The use of the Internet has become more and more popular among non-English speaker, especially for people in the countries like Japan, Germany, China, and Korea [21].

Companies have awakened to that trend of the Internet becoming multilingual and are developing country-specific and language-specific Web sites for individual markets [12]. Some examples of these corporations are IBM, Gap, McDonald, VISA, and Coca-Cola. International firms such as Sony, Mercedes Benz, Kodak are very international focused as well [17]. For global businesses to be successful, they have to become multilingual [2]. The localization of the Web culturally, geographically, and linguistically is a trend that will continue [20]. The result will be non-English speakers occupying more than half of all Web traffic in

their respective native languages [21]. English will NOT be the online lingua franca for long [1].

4. The Multilingual Internet

The Internet is moving to multilingual web sites. This is due to web page text written in the native language. The result of having the Web sites written in the users' native language brings a sense of familiarity that helps bridge the language gap [17]. To facilitate this, a new coding scheme for the user interface had to be developed.

ASCII is the coding scheme of the Internet. ASCII, American Standard Code for Information Interchange, was defined by the American National Standards Institute (ANSI) in the United States and by the International Organization for Standardization (ISO) worldwide [10]. The character set consists of the English alphabet. To resolve this limited character set, Unicode was developed. It is a multi-byte character representations system for computers, and provides for the encoding/exchanging of all the text of the world's languages [3]. Unicode uses 16 bits, which means that it can represent more than 65,000 unique characters [23]. Unicode allows programmers to provide software that ordinary people can use in their native language" [3]. Unicode is a character coding system designed to provide the ability to display text of the world's languages on web pages.

Internet users want domain names in their native language instead of today's English-language domain names. Unicode can also be used for domain names. This would require no changes to the DNS protocol or servers [13]. Multilingual domain name is still under implementation because the languages other than English do not use ASCII-based characters. There are several organizations working toward solution to implement non-English domain names on the Web [17]. The current solution involves converting foreign language characters into Unicode. Then encoding them in ASCII for transmission over the Internet as is currently done. It creates a presentation layer to display domain names to end users in their native languages [28].

5. English as the Source Language for Multilingual Web Pages

A web page consists of two parts: text content and source code. Text content is the text displayed on the web page. Unicode is used here to display the different characters of the variety of world languages. Examples of source code components are XML, Java, JavaScript, and HTML. By dividing the web page into these two parts, a company can extract all language and cultural specific elements, making the source code language-neutral and easily tailored for a local language by adding translations of text messages, icons, and cultural data. Hence, the web page is easily made multilingual.

Such a web page is able to run with any localized text content, without requiring source code changes or recompilation. It is this source code that incorporates

English into the infrastructure. Web developers write English "source code" to support multilingual web page "content." When you view the source code for any multilingual web page, the HTML, XML, Java, and JavaScript are in English.

Such software allows web pages for multilingual users to display the text in the correct local language and character set. For example, XML uses Unicode by default, which allows end users to encode text in their preferred language and character set [18].

Because ASCII is used by the Internet, the domain names are also in English. As stated earlier, the user interface will use Unicode, making domain name services multilingual. However, as noted, the Unicode domain names from the user interface will be changed to the English version of ASCII.

Considering the source code and domain names, English will continue to remain the most widely used language in the technical arena of the Internet.

6. English as the Global Language for Internet and other Information Systems Professionals

The language used at international Internet conferences and by international Internet journals is English. Table 1 shows some of the Internet related conferences for the year 2002. Research papers, Proceedings, official web site, and/or language of these conferences used English, and in some cases, along with the host country's native language. Table 2 shows some of the international journals relating to the Internet that are also in English.

7. Implications for Governments and Global Businesses

What appears to be developing on the Internet, as far as spoken languages are concern, is a distinction between local and global languages. The Internet will be multilingual due cultural pride and that people are more comfortable with e-commerce when the product is in their native language [24]. Local web interfaces will be in the local language. This includes operating systems and web development tools. Sun's Solaris 7 operating system is a fully internationalized. It uses a single global binary that allows for the use of different language versions [19]. Microsoft, IBM, and Sun are also developing their software systems in multi-languages [9].

When dealing on a global scale, there has to be a common language for all to understand. English is the language of international commerce/business and science [8][15][27]. European businesses are making English their official language. Companies are looking at the whole continent as their domestic market. [8]. English is becoming the common language for a multilingual continent. It appears that global developers and global users of the web are following this trend. To be able to

function on a Global/International scale via the Internet, Internet professionals and business executives will have English as a second language.

Table 1. Internet related conferences for the year 2002

15th Bled Electronic Commerce Conference eReality: Constructing the eEconomy Bled, Slovenia , June 17 - 19, 2002.
7 th Association Information Management Conference to be held in Hammamet, Tunisia, May 30 to June 1, 2002.
2 nd International Conference on Electronic Business (ICEB 2002), Taipei, Taiwan, Dec 10-13, 2002.
The 7th Asia Pacific Decision Sciences Institute (APDSI) Annual Meeting , Bangkok, Thailand, 24-27 July, 2002.
The 3rd International Conference On Web Information Systems Engineering (WISE 2002): 12-14th Dec 2002, Singapore.
The International Conference on Information Technology, Communications & Development , Kathmandu, Nepal, 1-3 December 2002.
11th European Conference on Information Systems , Naples, Italy, 19-21 June, 2003.
The First International Conference on Information and Management Sciences , Xi'An, China, 24-28 May, 2002.
Internet & Investments Forum , St. Petersburg, Russia, 25-26 April 2002
Second International Workshop on Electronic Commerce (WELCOM'01) , Heidelberg, Germany, 16-17 November, 2001.
Global IPv6 Summit in Korea July 11-12, 2002, Soul, Korea.

Table 2. International Journals relating to the Internet

Internet Research, MCB University Press, UK
International Journal of Electronic Business, Inderscience Publications, UK
International Journal of Internet and Enterprise Management, Inderscience Pub., UK
Journal of Global Information Management, Idea Group Publishing, USA
The International Journal of Information Systems Applications, Elsevier Science, USA

The impact is that computer curriculum in non-English speaking countries will have to incorporate English. Global businesses will have to provide English training to their international computer personnel.

8. Conclusion

The majority of Web page content will be in languages other than English [11]. Evidence of this growing trend can be seen through the emergence of the tools such as translation services, multilingual domain names, Unicode, and the use of icons [17]. It is clear that Web pages are becoming multilingual for users and business customers. However, the global infrastructure of the Internet will still in English because of the need for global standards in programming and technical communication. English will be the global language for Internet developers and global businesses professionals.

References

- [1] Dornan, A. (1999). The Power of Babel. *Data Communications*, 28(12) p.14.
- [2] Baker, S. (2001). A net Not Made in America. *Business Week*, 3725:124.
- [3] Burger, T. W. (2001). Linux Unicode programming: How to incorporate and utilize Unicode for foreign language support. <http://www-106.ibm.com/developerwork/linux/library/1-linumi.html>. Downloaded on 6/11/2002.
- [4] Castelluccio, M. (1999) Hey, Can Anybody Read This? *Management Accounting*, 81(1) July, pp.63-64.
- [5] Chapin, D. (2000). Going Global via the Web? Don't trip. *McTechnology Marketing Intelligence*, 20(11), pp. 54-58.
- [6] Crispin, S. W. (2000). Local Lingo. *Far Eastern Economic Review*, 163(18), p 43.
- [7] Dornan, A. (1999). The Power of Babel. *Data Communications*, 28(12), pp. 14.
- [8] Fox, J. (2000). The Triumph of English. *Fortune*, 142(6), pp. 209-212.
- [9] Gantz, J. (1998). Coming soon: Language Barriers. *Computerworld*, 32(5), p. 33.
- [10] Gilbert, H. (1999) Understanding Data Communications, 6th ed. New Riders, Indianapolis.
- [11] Glaeser, C. (2000) Local Heroes. *Bank Systems & Technology*, 37(10) October, S5, S17.
- [12] Gray, R. (2000). Make the most of the Local Difference. *Marketing*, 27-28.
- [13] Marsan, C. D. (2001). Patent flap slows multilingual domain name plan. *Network World*, 18(13), pp. 9,30.
- [14] MacLeod, M. (2000). Language Barriers. *Supply Management* 5(14), pp. 37-38.
- [15] McCune, J. C. (1999). English written here. *Management Review*, 88(2), p. 12.
- [16] Oz, E. (2000). *Management Information Systems*, 2nd Ed. Course Technology - Thomas Learning, Cambridge, Ma.
- [17] Padunchwit, P. and White, G. (2001). From English to Multilingual on the Internet. Proceedings of the 33rd Annual Conference of the Southwest Decision Sciences Conference, St. Louis, Mo. November, 2001.
- [18] Schloss, W. (2000). Ten best bets for XML applications. <http://www-106.ibm.com/developerwork/library/tenxmlapps/>. Accessed on 6/11/2002.
- [19] Sun Microsystems (1998). Solaris 7 Globalization

Technology Overview: Product Brief. Sun Microsystems, Inc. Palo Alto, California.

[20] Taylor, C. (2000). The Language of the Web. *Far Eastern Economic Review*, 163(49), pp. 68-69.

[21] Tweney, D. (1999) Increasingly Global, the Web Challenges U.S.-Based Companies. *Infoworld*, 21(29) July, p.52.

[22] Ulfelder, S. (2000). All the Web's a Stage. *CIO*, 14(1), pp. 132-142.

[23] Unicode, Inc. (2001) What is Unicode, Available from: <http://unicode.org/unicode/standard/WhatIsUnicode.html>. Accessed 29 April, 2001.

[24] Wah, L. (1998). Online subscribers booming among non-English speakers. *Management Review*, 87(6), p.7.

[25] Wilhelm, K. (2001). Tongue-tied on the Net. *Far Eastern*

Economic Review, 164(7), pp.38-40.

[26] Wonnacott, L. (2001). Going Global may bring New Opportunities for Existing Customers. *Infoworld*, 23(14), pp 58.

[27] Woodward, N. H. (1999). Do you speak Internet? *HR Magazine*, 44(4), pp. s12-s16.

[28] Yamada, M. (2000) Tackling the Web's Language Barrier, Available from:

<http://www.cnn.com/2000/tech/computing/07/14/icann.language.idg/index.html>. Access 20 April, 2001.

[29] Yunker, J. (2000). Going Global. *Pharmaceutical Executive*, 20(7), pp. 138-146.